

FLUCTUATIONS OF MATRIX ENTRIES OF ANALYTIC FUNCTIONS OF NON-HERMITIAN RANDOM MATRICES

SEAN O'ROURKE

ABSTRACT. Consider an $n \times n$ non-Hermitian random matrix M_n whose entries are independent real random variables. Under suitable conditions on the entries, we study the fluctuations of the entries of $f(M_n)$ as n tends to infinity, where f is analytic on an appropriate domain. This extends the results in [19], [20], and [23] from symmetric random matrices to the non-Hermitian case.

1. INTRODUCTION AND MAIN RESULTS

There have been a number of results recently concerning matrix entries of functions of random matrices. The results so far have dealt with symmetric (Hermitian) random matrix ensembles. That is, for a $n \times n$ real symmetric (Hermitian) random matrix, M_n , one studies the entries of the matrix $f(M_n)$ where f is a regular test function.

For a symmetric (Hermitian) matrix M_n , we define $f(M_n)$ using functional calculus. That is, if M_n is diagonalized as $M_n = V^* DV$ where V is an orthogonal (unitary) matrix and D is diagonal, then $f(M_n)$ is constructed as $f(M_n) = V^* f(D)V$ where $f(D)$ is again a diagonal matrix with the entries $f(D_{ii})$ on the diagonal.

In [16], Lytova and Pastur consider the case where M_n is drawn from the Gaussian Orthogonal Ensemble (GOE) or Gaussian Unitary Ensemble (GUE). The more general case where M_n is a symmetric (Hermitian) Wigner matrix with i.i.d. entries (not necessarily Gaussian) was studied by Pizzo, Renfrew, and Soshnikov in [23] and by Pastur and Lytova in [22].

The results of [23] and [22] are extended in [19] to the case where M_n is a Wigner matrix with independent but not necessarily identically distributed entries. The authors require that the off-diagonal entries have uniformly bounded fourth moments, diagonal entries have uniformly bounded second moments, and certain Lindeberg type conditions for both the off-diagonal entries and the diagonal entries are satisfied. The test function $f(x)$ is assumed to satisfy

$$\int_{\mathbb{R}} (1 + 2|k|)^{2s} |\hat{f}(k)|^2 dk < \infty,$$

for some $s > 3$ where \hat{f} is the Fourier transform

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ikx} f(x) dx.$$

Under these conditions it is shown in [19] that

$$(1) \quad \sqrt{n} \left(f \left(\frac{M_n}{\sqrt{n}} \right)_{ij} - \mathbb{E} f \left(\frac{M_n}{\sqrt{n}} \right)_{ij} \right) - \alpha(f)(M_n)_{ij} \longrightarrow N(0, d^2(f)),$$

as $n \rightarrow \infty$, where $\alpha(f)$ and $d^2(f)$ are given constants depending on the function f .

Recently, similar results were obtained in [20] for the sample covariance case where $M_n = A_n A_N^*$ and A_n is an $n \times N$ random matrix with independent entries under analogous moment and Lindeberg type conditions on the entries of A_n .

Instead of considering a specific entry of $f(M_n)$, Bai and Pan in [6] study the fluctuations of $u^* f(M_n) u$ when M_n is a Wigner matrix whose entries satisfy some moment assumptions, f is an analytic function, and u is a delocalized unit vector.

In this paper, we consider non-Hermitian random matrices with independent entries. Given a $n \times n$ random matrix M_n , we define the empirical spectral distribution (ESD), F^{M_n} , of M_n to be

$$F^{M_n}(x, y) := \frac{1}{n} \# \{1 \leq j \leq n : \operatorname{Re}(\lambda_j) \leq x, \operatorname{Im}(\lambda_j) \leq y\},$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of M_n .

The ESD for non-Hermitian random matrices with i.i.d. Gaussian entries was first studied by Mehta [17]. In particular, in the case where the entries of M_n are i.i.d. standard complex normal random variables, the ESD of $\frac{1}{\sqrt{n}} M_n$ converges, as $n \rightarrow \infty$, to the circular law F_c given by

$$F_c(x, y) := \frac{1}{\pi} \operatorname{mes} \left(|z| \leq 1 : \operatorname{Re}(z) \leq x, \operatorname{Im}(z) \leq y \right).$$

In other words, F_c is the uniform distribution over the unit disk in the complex plane.

Mehta used the joint density function of the eigenvalues of $\frac{1}{\sqrt{n}} M_n$ which was derived by Ginibre [11]. The real Gaussian case was studied by Edelman in [10]. For the general (non-Gaussian) case when there is no formula, the problem appears much more difficult. Important results were obtained by Girko [12, 13], Bai [2, 4], and more recently by Götze and Tikhomirov [14], Pan and Zhou [21], and Tao and Vu [26]. These results confirm the same limiting law under some moment or smoothness assumptions on the distribution of the entries. Recently, Tao and Vu (appendix by Krishnapur) were able to remove all these additional assumptions, establishing the law under the first two moments [27].

Theorem 1 (Circular law for non-Hermitian i.i.d. matrices; [27]). *Assume that the entries of the $n \times n$ matrix M_n are i.i.d. copies of a complex random variable of mean zero and variance one. Then the ESD of the matrix $\frac{1}{\sqrt{n}} M_n$ converges almost surely to F_c as $n \rightarrow \infty$.*

Theorem 1 also holds for random matrices with independent but not necessarily identically distributed entries, provided one has some uniform control of the entries (see Appendix C of [27]).

When M_n is a non-Hermitian matrix and f is analytic, we define $f(M_n)$ as

$$f(M_n) = \sum_{k=0}^{\infty} f^{(k)}(0) \frac{M_n^k}{k!}$$

provided

$$\sum_{k=0}^{\infty} |f^{(k)}(0)| \frac{\|M_n^k\|}{k!} < \infty,$$

where $\|\cdot\|$ denotes the spectral norm.

In [25], Rider and Silverstein study linear eigenvalue statistics of non-Hermitian random matrices with i.i.d. entries. Let f be analytic in a neighborhood of the disk $\{z \in \mathbb{C} : |z| \leq 4\}$. It is shown in [25] that

$$(2) \quad \text{tr} \left[f \left(\frac{1}{\sqrt{n}} M_n \right) \right] - nf(0)$$

converges to a mean zero complex Gaussian as the size of the matrix tends to infinity, provided the distribution of the complex entries of M_n satisfy some moment and smoothness assumptions. The real case was studied in [18] under a different set of assumptions on the entries of M_n in the case when f is a polynomial.

The goal of this paper is to prove the analogue of (1) in the case that M_n is a non-Hermitian random matrix with independent entries. In particular, we consider the following ensemble of non-Hermitian random matrices.

Definition 2. For each $n \geq 1$, let $M_n = (m_{nij})$ be a $n \times n$ real random matrix. The sequence of random matrices $\{M_n\}_{n \geq 1}$ is said to satisfy condition **C0** if the following conditions hold:

- (i) For each $n \geq 1$, $\{m_{nij} : 1 \leq i, j \leq n\}$ is a collection of independent real random variables with mean zero and variance one,
- (ii) $\sup_{n,i,j} \mathbb{E}|m_{nij}|^4 < \infty$,
- (iii) For all $\epsilon > 0$,

$$\frac{1}{n^2} \sum_{i,j=1}^n \mathbb{E}|m_{nij}|^4 \mathbf{1}_{\{|m_{nij}| > \epsilon \sqrt{n}\}} \longrightarrow 0$$

as $n \rightarrow \infty$.

Remark 3. For each $n \geq 1$, let M_n be an $n \times n$ matrix whose entries are i.i.d. copies of a real random variable ξ . Then the sequence $\{M_n\}_{n \geq 1}$ satisfies condition **C0** if ξ has mean zero, unit variance, and $\mathbb{E}[\xi^4] = m_4 < \infty$.

For each $n \geq 1$, let $M_n = (m_{nij})$ be a $n \times n$ real random matrix and assume the sequence of random matrices $\{M_n\}_{n \geq 1}$ satisfies condition **C0**. Let f be analytic in a region that contains the disk $\{z \in \mathbb{C} : |z| \leq 2 + \epsilon\}$ for some $\epsilon > 0$. In our main result below, we show that for any fixed indices i and j , the entry

$$(3) \quad f \left(\frac{1}{\sqrt{n}} M_n \right)_{ij} - f(0)\delta_{ij} - f'(0) \frac{1}{\sqrt{n}} m_{nij}$$

converges in distribution to a complex Gaussian random variable as n tends to infinity, where δ_{ij} denotes the Kronecker delta. The appearance of $f(0)$ is not surprising considering (2) and since $f(0)$ is the average of f against the circular law. That is,

$$\frac{1}{\pi} \int_{\{z:|z|\leq 1\}} f(z) d^2 z = f(0)$$

because f is analytic. The $f'(0)$ term in (3) is analogous to the $\alpha(f)$ term which appears in (1).

Theorem 4. For each $n \geq 1$, let M_n be an $n \times n$ real random matrix and assume the sequence $\{M_n\}_{n \geq 1}$ satisfies condition **C0**. Fix $l \geq 1$ and assume that for each $1 \leq i \leq l$ and every $\delta > 0$,

$$(4) \quad \frac{1}{n} \sum_{j=1}^n \left\{ \mathbb{E} \left[m_{nij}^4 \mathbf{1}_{\{|m_{nij}| > \delta n^{1/4}\}} \right] + \mathbb{E} \left[m_{nji}^4 \mathbf{1}_{\{|m_{nji}| > \delta n^{1/4}\}} \right] \right\} \longrightarrow 0$$

as $n \rightarrow \infty$. Then the following holds:

- (i) if f is analytic in a region that contains the disk $\{z \in \mathbb{C} : |z| \leq 2 + \epsilon\}$ for some $\epsilon > 0$, the normalized matrix entries

$$\left\{ \sqrt{n} \left(f \left(\frac{1}{\sqrt{n}} M_n \right)_{ij} - f(0)\delta_{ij} - f'(0) \frac{1}{\sqrt{n}} m_{nij} \right) : 1 \leq i, j \leq l \right\}$$

are independent in the limit $n \rightarrow \infty$,

- (ii) if f_1, \dots, f_m are analytic in a region that contains the disk $\{z \in \mathbb{C} : |z| \leq 2 + \epsilon\}$ for some $\epsilon > 0$ and $1 \leq i, j \leq l$ are fixed, the random m -vector

$$\sqrt{n} \left(f_k \left(\frac{1}{\sqrt{n}} M_n \right)_{ij} - f_k(0)\delta_{ij} - f'_k(0) \frac{1}{\sqrt{n}} m_{nij} \right)_{k=1}^m$$

converges in distribution to the mean zero multivariate complex Gaussian $(Z(f_1), \dots, Z(f_m))$ as $n \rightarrow \infty$ defined by

$$(5) \quad \mathbb{E}[Z(f_p) \overline{Z(f_q)}] = \sum_{r=2}^{\infty} \frac{f_p^{(r)}(0)}{r!} \frac{\overline{f_q^{(r)}(0)}}{r!}.$$

Remark 5. We draw the reader's attention to one special case of Theorem 4. Consider the monomial $f(z) = z^t$ for some integer $t > 1$. In this case, the limiting distribution $Z(f)$ is standard Gaussian. Moreover, if $f(z) = z^t$, $g(z) = z^s$ for integers $s, t > 1$ with $s \neq t$, then (5) implies that the limiting distributions $Z(f)$ and $Z(g)$ are independent.

Theorem 4 is a corollary of Theorem 6 below in which we study the entries of the resolvent

$$\left(zI_n - \frac{1}{\sqrt{n}} M_n \right)^{-1} = \left(z - \frac{1}{\sqrt{n}} M_n \right)^{-1},$$

where I_n is the identity matrix of order n . The result can be extended from the resolvent case to arbitrary analytic functions by using Cauchy's integral formula.

Below and throughout the paper, we use $\sqrt{-1}$ for the imaginary unit and reserve i for an index.

Theorem 6. For each $n \geq 1$, let M_n be an $n \times n$ real random matrix and assume the sequence $\{M_n\}_{n \geq 1}$ satisfies condition **C0**. Fix $l \geq 1$ and assume that for each $1 \leq i \leq l$ and all $\delta > 0$ (4) holds as $n \rightarrow \infty$. Define the $l \times l$ matrix

$$(6) \quad Y_n(z) := \sqrt{n} \left[\left(z - \frac{1}{\sqrt{n}} M_n \right)_{ij}^{-1} - \frac{1}{z} \delta_{ij} - \frac{m_{nij}}{z^2 \sqrt{n}} \right]_{i,j=1}^l.$$

Let $D \subset \mathbb{C}$ be any compact set outside the disk $\{z \in \mathbb{C} : |z| \leq 2 + \epsilon\}$ for some $\epsilon > 0$. Define

$$S_n(z) := \sum_{i,j=1}^l \left[\alpha_{ij} \left(\frac{Y_n(z)_{ij} + \overline{Y_n(z)_{ij}}}{2} \right) + \beta_{ij} \left(\frac{Y_n(z)_{ij} - \overline{Y_n(z)_{ij}}}{2\sqrt{-1}} \right) \right]$$

where $\alpha_{ij}, \beta_{ij} \in \mathbb{C}$ for $1 \leq i, j \leq l$. Then $\{S_n\}_{n \geq 1}$ is tight in the space of continuous functions on D and converges weakly to the complex Gaussian process

$$S(z) := \sum_{i,j=1}^l \left[\alpha_{ij} \left(\frac{Y(z)_{ij} + \overline{Y(z)_{ij}}}{2} \right) + \beta_{ij} \left(\frac{Y(z)_{ij} - \overline{Y(z)_{ij}}}{2\sqrt{-1}} \right) \right]$$

where $Y(z) = \{Y(z)_{ij}\}_{i,j=1}^l$ is the matrix valued complex Gaussian process defined by

$$\mathbb{E}[Y(z)_{ij}\overline{Y(w)}_{ij}] = \frac{1}{z^2\bar{w}^2} \frac{1}{z\bar{w}-1},$$

and $\mathbb{E}[Y(z)_{ij}] = 0$.

In addition, for any finite $s \geq 1$, the entries $Y_n(z_k)_{i_k, j_k}$, $1 \leq i_k, j_k \leq l$, $1 \leq k \leq s$, $z_1, \dots, z_s \in D$ are independent provided $(i_1, j_1), (i_2, j_2), \dots, (i_s, j_s)$ are distinct.

The fact that Theorem 6 is stated for D outside the disk $\{z \in \mathbb{C} : |z| \leq 2 + \epsilon\}$ (and Theorem 4 is stated for functions analytic in a region containing $\{z \in \mathbb{C} : |z| \leq 2 + \epsilon\}$) requires some explanation. For simplicity, assume for each $n \geq 1$, M_n is a $n \times n$ real random matrix whose entries are i.i.d. copies of a random variable with mean zero, unit variance, and finite fourth moment. Since the spectral radius of $\frac{1}{\sqrt{n}}M_n$ converges to 1 in probability as n tends to infinity [4, Theorem 5.18], it is natural to conjecture that Theorem 6 should hold for D outside the disk $\{z \in \mathbb{C} : |z| \leq 1 + \epsilon\}$ (and Theorem 4 should hold for f analytic in a region containing $\{z \in \mathbb{C} : |z| \leq 1 + \epsilon\}$).

In order to prove Theorem 6, we need to ensure that $zI_n - \frac{1}{\sqrt{n}}M_n$ is invertible and obtain control of the resolvent. Since the spectral radius converges to 1, it is natural to work on the event that the spectral radius is less than $1 + \epsilon/2$. However, in the proof of Theorem 6, we must also consider the matrix $M_{n,l}$, which is constructed from the matrix M_n by removing the first l columns and l rows. The problem that arises is that control of the spectral radius of M_n does not imply that we have similar control of the spectral radius of $M_{n,l}$. On the other hand, $\|M_{n,l}\| \leq \|M_n\|$. Thus we must deal with the spectral norm instead of the spectral radius. Since $\frac{1}{\sqrt{n}}\|M_n\|$ converges in probability to 2 as $n \rightarrow \infty$ (see Lemma 8 below), we have uniform control of the resolvent if we take $|z| > 2$. This statement is made precise in Proposition 11.

The paper is organized as follows. In Section 2, we state our preliminary tools and prove Theorem 4. The proof of Theorem 6 is divided into two parts and presented in Sections 3 and 4.

2. OVERVIEW AND PRELIMINARY TOOLS

In this section, we introduce some standard notation and present our preliminary tools.

Throughout this paper, n is an asymptotic parameter going to infinity. We use $o(1)$ to denote any quantity that is bounded in magnitude by an expression that converges to zero as n tends to infinity. We use $X = O(Y)$ to denote the estimate $|X| \leq CY$ where the implied constant C is independent of n . We use C throughout the paper to denote some positive constant which does not depend on n that may change from line to line.

We note that by condition (iii) from Definition 2, it follows that there exists a sequence $\epsilon_n \rightarrow 0$ such that

$$\frac{1}{\epsilon_n^4 n^2} \sum_{i,j} \mathbb{E}|m_{nij}|^4 \mathbf{1}_{\{|m_{nij}| > \epsilon_n \sqrt{n}\}} \longrightarrow 0$$

as $n \rightarrow \infty$.

We then have the following standard truncation lemma which can be found in [20].

Lemma 7. Assume that for each $n \geq 1$, M_n is an $n \times n$ real random matrix and the sequence $\{M_n\}_{n \geq 1}$ satisfies condition **C0**. Fix $l \geq 1$ and write $[l] = \{1, 2, \dots, l\}$. Assume that for each $i \in [l]$ and for every $\delta > 0$, (4) holds as $n \rightarrow \infty$. Then there exists a sequence $\{\tilde{M}_n\}_{n \geq 1}$ where \tilde{M}_n is a $n \times n$ matrix with independent entries and there exists a sequence ϵ_n which tends to zero as n tends to infinity such that

- (i) the entries $(\tilde{M}_n)_{jk}$ have mean zero and variance one,
- (ii) $\sup_{j,k \notin [l]} |(\tilde{M}_n)_{jk}| \leq \epsilon_n \sqrt{n}$,
- (iii) $\sup_{1 \leq j \leq n, i \in [l]} |(\tilde{M}_n)_{ij}| \leq \epsilon_n n^{1/4}$,
- (iv) $\sup_{1 \leq j \leq n, i \in [l]} |(\tilde{M}_n)_{ji}| \leq \epsilon_n n^{1/4}$,
- (v) $\sup_{n,j,k} \mathbb{E}|(\tilde{M}_n)_{jk}|^4 < \infty$,
- (vi) $\mathbb{P}(M_n \neq \tilde{M}_n) \rightarrow 0$ as $n \rightarrow \infty$.

Using [4, Theorem 5.9] and Lemma 7, we obtain the following bound on the spectral norm of M_n .

Lemma 8. For each $n \geq 1$, let M_n be an $n \times n$ real random matrix and assume the sequence $\{M_n\}_{n \geq 1}$ satisfies condition **C0**. Then

$$\limsup_{n \rightarrow \infty} \left\| \frac{M_n}{\sqrt{n}} \right\| \leq 2$$

in probability.

Assume Theorem 6 for the moment. We are now ready prove Theorem 4.

Proof of Theorem 4. Fix $\epsilon > 0$ and let γ be the circle of radius $2 + \epsilon$ centered around the origin. Let D be a compact set outside the disk $\{z \in \mathbb{C} : |z| \leq 2 + \epsilon/2\}$ that contains γ . Suppose f is analytic in a disk about the origin that contains the contour γ . By Lemma 8 and Cauchy's integral formula, we have, with probability going to one, that

$$\sqrt{n} \left(f \left(\frac{1}{\sqrt{n}} M_n \right)_{ij} - f(0) \delta_{ij} - f'(0) \frac{1}{\sqrt{n}} m_{nij} \right) = \frac{1}{2\pi\sqrt{-1}} \int_{\gamma} f(z) Y_n(z)_{ij} dz$$

for $1 \leq i, j \leq l$. To study the joint distribution of the l^2 entries, we use the Cramer-Wold device and study random variables of the form

$$(7) \quad \frac{1}{2\pi\sqrt{-1}} \int_{\gamma} (f_1(z) + \dots + f_k(z)) S_n(z) dz,$$

where f_1, \dots, f_k are analytic in a region about the origin that contains the contour γ . By Theorem 6, $S_n(z)$ converges weakly to $S(z)$ in the space of continuous functions on D . Therefore the random variable in (7) converges in distribution to a Gaussian random variable with mean zero as n tends to infinity. We now compute the limiting covariances. Since $S(z)$ is a linear combination of the entries of $Y(z)$, it suffices to study the entries of $Y(z)$. For functions f and g analytic in a region that contains the contour γ , the limiting covariance is given by

$$\mathbb{E}[Z(f) \overline{Z(g)}] = \mathbb{E} \left[\frac{1}{2\pi\sqrt{-1}} \int_{\gamma} f(z) Y(z)_{ij} dz \overline{\frac{1}{2\pi\sqrt{-1}} \int_{\gamma} g(w) Y(w)_{ij} dw} \right].$$

Using Fubini's theorem, we can rewrite the limiting covariance as

$$\begin{aligned}\mathbb{E}[Z(f)\overline{Z(g)}] &= \left(\frac{1}{2\pi}\right)^2 \int_{\gamma} \int_{\gamma} f(z) \overline{g(w)} \mathbb{E}[Y(z)_{ij} \overline{Y(w)_{ij}}] dz d\bar{w} \\ &= \left(\frac{1}{2\pi}\right)^2 \int_{\gamma} \int_{\gamma} \frac{f(z)}{z^2} \frac{\overline{g(w)}}{w^2} \frac{1}{z\bar{w} - 1} dz d\bar{w} \\ &= \sum_{r=2}^{\infty} \frac{f^{(r)}(0)}{r!} \frac{\overline{g^{(r)}(0)}}{r!}.\end{aligned}$$

Since $\mathbb{E}[Y(z)_{i_1 j_1} Y(w)_{i_2 j_2}] = 0$ for $(i_1, j_1) \neq (i_2, j_2)$ by Theorem 6 and since the joint limiting distribution is Gaussian, it follows that the normalized entries

$$\left\{ \sqrt{n} \left(f \left(\frac{1}{\sqrt{n}} M_n \right)_{ij} - f(0) \delta_{ij} - f'(0) \frac{1}{\sqrt{n}} m_{nij} \right) : 1 \leq i, j \leq l \right\}$$

are independent in the limit $n \rightarrow \infty$. The proof of Theorem 4 is complete. \square

We now turn our attention to proving Theorem 6. We will continue to use the notation $\|B\|$ to denote the spectral norm of the matrix B . We will also make use of the following lemmas.

Lemma 9. *Let \mathbf{u}_n and \mathbf{v}_n be independent n -vectors whose entries are independent random variables with mean zero and variance one. Let B_n be an independent $n \times n$ random matrix where $\|B\| \leq a$ for all n . Then*

$$\mathbb{E} \left| \frac{1}{\sqrt{n}} \mathbf{u}_n^T B \mathbf{v}_n \right|^2 = \frac{1}{n} \mathbb{E} \text{Tr}(BB^T) \leq a^2.$$

Proof. The inequality is trivial and follows from the assumption $\|B\| \leq a$. We denote the entries of $\mathbf{u}_n = (u_j)_{j=1}^n$, $\mathbf{v}_n = (v_k)_{k=1}^n$, and $B = (B_{ij})_{i,j=1}^n$. Then by the independence of B , \mathbf{u}_n , and \mathbf{v}_n , we have that

$$\begin{aligned}\mathbb{E} \left| \frac{1}{\sqrt{n}} \mathbf{u}_n^T B \mathbf{v}_n \right|^2 &= \frac{1}{n} \sum_{j,k,l,m=1}^n \mathbb{E} u_j B_{jk} v_k u_l B_{lm} v_m \\ &= \frac{1}{n} \sum_{j,k} \mathbb{E} B_{jk}^2 = \frac{1}{n} \mathbb{E} \text{Tr}(BB^T).\end{aligned}$$

\square

Lemma 10. *For each $n \geq 1$, let $\mathbf{x}_n = (x_{ni})_{i=1}^n$ be a real n -vector and let $B_n = (b_{nij})_{i,j=1}^n$ be an independent $n \times n$ random matrix. Assume that*

- (i) *The entries of \mathbf{x}_n are independent and have mean zero and variance 1,*
- (ii) $\sup_{n,i} \mathbb{E}|x_{ni}|^4 \leq m_4 < \infty$,
- (iii) *There exists a sequence $\epsilon_n \rightarrow 0$ such that $|x_{ni}| \leq \epsilon_n \sqrt{n}$ for all $n \geq 1$ and $1 \leq i \leq n$,*
- (iv) *There exists a constant a (not depending on n) such that $\|B\| \leq a$.*

Then there exists an absolute constant $C > 0$ such that

- (i) $\mathbb{E} \left| \frac{1}{n} \mathbf{x}_n^T \mathbf{x}_n - 1 \right|^4 \leq C \left[\frac{m_4 \epsilon_n}{n} + \frac{m_4^2}{n^2} \right]$,
- (ii) $\mathbb{E} \left| \frac{1}{n} \mathbf{x}_n^T B_n \mathbf{x}_n - \frac{1}{n} \text{tr}(B_n) \right|^4 \leq C a^4 m_4 \left(\frac{1}{n^2} + \frac{\epsilon_n}{n} \right)$,
- (iii) *For any fixed index k , $\mathbb{E} \left| \frac{1}{\sqrt{n}} e_k^T B_n \mathbf{x}_n \right|^4 \leq C \left[a^4 \frac{m_4 + 1}{n^2} \right]$,*

(iv) For any fixed index k , $\mathbb{E}|\frac{1}{\sqrt{n}}\mathbf{x}_n^T B_n e_k|^4 \leq C [a^4 \frac{m_4+1}{n^2}]$.

Proof. To prove (i), we note that

$$\mathbb{E}|\frac{1}{n}\mathbf{x}_n^T \mathbf{x}_n - 1|^4 = \frac{1}{n^4} \sum_{j_1, j_2, j_3, j_4=1}^n \mathbb{E}[(x_{nj_1}^2 - 1)(x_{nj_2}^2 - 1)(x_{nj_3}^2 - 1)(x_{nj_4}^2 - 1)].$$

By the independence of the entries of \mathbf{x}_n , $\mathbb{E}[(x_{nj_1}^2 - 1)(x_{nj_2}^2 - 1)(x_{nj_3}^2 - 1)(x_{nj_4}^2 - 1)] = 0$ when any j_l is distinct from the other three indices. So we have two cases that give non-vanishing expectation:

(1) When $j_1 = j_2 = j_3 = j_4$, we have that

$$\frac{1}{n^4} \sum_{j=1}^n \mathbb{E}[(x_{nj}^2 - 1)^4] \leq \frac{1}{n^4} nm_4 (\epsilon_n \sqrt{n})^4 \leq \frac{m_4}{n} \epsilon_n.$$

(2) For $j_1 = j_l$ and $j_s = j_t$ with $j_s \neq j_1$, there are 3 arrangements of such pairings. For each pairing, we have that

$$\frac{1}{n^4} \sum_{j_1 \neq j_2} \mathbb{E}[(x_{nj_1}^2 - 1)^2 (x_{nj_2}^2 - 1)^2] \leq \frac{1}{n^4} n^2 m_4^2$$

Thus, we have that

$$\mathbb{E}|\frac{1}{n}\mathbf{x}_n^T \mathbf{x}_n - 1|^4 \leq C \left[\frac{m_4 \epsilon_n}{n} + \frac{m_4^2}{n^2} \right].$$

Statement (ii) follows from Lemma B.26 in [4] (taking $p = 4$) by conditioning on the entries of B_n since B_n and \mathbf{x}_n are independent.

To prove statement (iii), we write

$$\mathbb{E}|\frac{1}{\sqrt{n}}e_k^T B_n \mathbf{x}_n|^4 = \frac{1}{n^2} \sum_{j_1, j_2, j_3, j_4=1}^n \mathbb{E}[B_{nkj_1} x_{nj_1} B_{nkj_2} x_{nj_2} \overline{B_{nkj_3}} x_{nj_3} \overline{B_{nkj_4}} x_{nj_4}].$$

As in the proof of statement (i), we use the independence of the entries and consider two distinct cases when the expectation of the summand is non-zero. The cases are:

(1) When $j_1 = j_2 = j_3 = j_4$, we have that

$$\frac{1}{n^2} \sum_{j=1}^n \mathbb{E}[|B_{nkj}|^4 x_{nj}^4] \leq \frac{m_4}{n^2} a^2 (B_n B_n^*)_{kk} \leq \frac{m_4 a^4}{n^2}.$$

(2) For $j_1 = j_l$ and $j_s = j_t$ with $j_s \neq j_1$, there are 3 arrangements of such pairings. In each pairing we obtain

$$\frac{1}{n^2} \sum_{j_1 \neq j_2} \mathbb{E}[B_{nkj_1}^2 B_{nkj_2}^2 x_{nj_1}^2 x_{nj_2}^2] \leq \frac{1}{n^2} (B_n B_n^*)_{kk}^2 \leq \frac{a^4}{n^2}.$$

Combining the two bounds above completes the proof of statement (iii). The proof of statement (iv) follows the same argument. \square

For the remainder of the paper, we fix $l \geq 1$, $\epsilon > 0$, and a compact set $D \subset \mathbb{C}$ outside the disk $\{z \in \mathbb{C} : |z| \leq 2 + \epsilon\}$.

We write the matrix M_n as a block matrix in the following way: let X be the upper-left $l \times l$ block of M_n , let $M_{n,l}$ be the lower-right $(n-l) \times (n-l)$ block of

M_n , let ϕ be the lower-left $(n-l) \times l$ block of M_n , and let ψ be the upper-right $l \times (n-l)$ block. That is,

$$M_n = \begin{bmatrix} X & \psi \\ \phi & M_{n,l} \end{bmatrix}.$$

Since the entries of M_n are independent, $X, M_{n,l}, \phi$, and ψ are independent.

We write the resolvent of M_n as

$$R_n(z) := \left(z - \frac{1}{\sqrt{n}} M_n \right)^{-1}$$

and write the resolvent of $M_{n,l}$ as

$$R_{n,l}(z) := \left(z - \frac{1}{\sqrt{n}} M_{n,l} \right)^{-1}.$$

In order to work with $R_n(z)$ (or $R_{n,l}(z)$) we will need a priori control of the norm $\|R_n(z)\|$ ($\|R_{n,l}(z)\|$). Following Rider and Silverstein in [25], we define the event

$$\Omega_n = \left\{ M_n : \frac{1}{\sqrt{n}} \|M_n\| \leq \kappa \right\}$$

where we fix the value $2 < \kappa \leq 2 + \epsilon/2$. Then for $|z| \geq 2 + \epsilon$, there exists a constant $\mathcal{K} = \mathcal{K}(\kappa, \epsilon)$ such that

$$\sup_{|z| \geq 2 + \epsilon} \|R_n(z)\| \leq \mathcal{K} \text{ on the event } \Omega_n.$$

Indeed, for $\frac{1}{\sqrt{n}} \|M_n\| \leq \kappa \leq 2 + \epsilon/2$,

$$\|R_n(z)\| = \left\| \frac{1}{z} \left(I_n - \frac{1}{z\sqrt{n}} M_n \right)^{-1} \right\| \leq \frac{1}{|z|} \sum_{k=0}^{\infty} \left\| \frac{M_n}{z\sqrt{n}} \right\|^k \leq \sum_{k=0}^{\infty} \left(\frac{2 + \epsilon/2}{2 + \epsilon} \right)^k,$$

provided $|z| \geq 2 + \epsilon$.

We now assume both κ and \mathcal{K} are fixed and restate the above result in the following proposition.

Proposition 11. *If $\frac{1}{\sqrt{n}} \|M_n\| \leq \kappa$, then $\sup_{|z| \geq 2 + \epsilon} \|R_n(z)\| \leq \mathcal{K}$.*

By the contrapositive of Proposition 11, if $\sup_{z \in D} \|R_n(z)\| > \mathcal{K}$, then $\frac{1}{\sqrt{n}} \|M_n\| > \kappa$. Thus,

$$\mathbb{P} \left(\sup_{z \in D} \|R_n(z)\| > \mathcal{K} \right) \leq \mathbb{P}(\Omega_n^C) \longrightarrow 0$$

as $n \rightarrow \infty$ by Lemma 8.

Therefore, when we work with the resolvent $R_n(z)$, we will always work on the event that $\sup_{z \in D} \|R_n(z)\|$ is finite. For the remainder of the paper, we will assume $z \in D$.

We also define the event

$$\Omega_{n,l} = \left\{ M_n : \frac{1}{\sqrt{n}} \|M_{n,l}\| \leq \kappa \right\}.$$

In particular, if $\|M_n\| \leq \kappa$ then $\|M_{n,l}\| \leq \kappa$ and hence $\Omega_n \subset \Omega_{n,l}$. Equivalently, we write $\mathbf{1}_{\Omega_n} \leq \mathbf{1}_{\Omega_{n,l}}$ where $\mathbf{1}_\Omega$ denotes the indicator function of the event Ω .

We divide the proof of Theorem 6 into two parts. In Section 3, we prove the finite dimensional distributions of $S_n(z)$ converge to $S(z)$. We then show that $\{S_n\}_{n \geq 1}$ is tight in the space of continuous functions on D in Section 4.

3. CONVERGENCE OF THE FINITE DIMENSIONAL DISTRIBUTIONS

In this section, we show that the finite dimensional distributions of $S_n(z)$ converge to $S(z)$. To study the entries of the resolvent we apply the standard formula for the inverse of a partitioned matrix (see for instance [15]). In particular, we obtain

$$\begin{aligned} \left[\left(z - \frac{1}{\sqrt{n}} M_n \right)_{ij}^{-1} \right]_{i,j=1}^l \mathbf{1}_{\Omega_{n,l}} &= \left(z - \frac{1}{\sqrt{n}} X - \frac{1}{n} \psi R_{n,l}(z) \phi \right)^{-1} \mathbf{1}_{\Omega_{n,l}} \\ &= \frac{1}{z} \left(I_l - \left[\frac{1}{\sqrt{n}} X + \frac{1}{n} \psi R_{n,l}(z) \phi \right] \right)^{-1} \mathbf{1}_{\Omega_{n,l}} \end{aligned}$$

with probability going to one.

It will follow from the discussion below and from Lemma 9 that

$$\|X\| + \|n^{-1/2} \psi R_{n,l}(z) \phi\| \mathbf{1}_{\Omega_{n,l}}$$

is bounded in probability. Thus, we have that

$$Y_n(z) \mathbf{1}_{\Omega_{n,l}} = \frac{\sqrt{n}}{z^2} \left[\frac{1}{n} \psi R_{n,l}(z) \phi \right] \mathbf{1}_{\Omega_{n,l}} + o(1)$$

in probability.

By Proposition 11,

$$\mathbb{P} \left(\|Y_n(z)\| \mathbf{1}_{\Omega_{n,l}^C} > \eta \right) \leq \mathbb{P}(\Omega_{n,l}^C) + \mathbb{P} \left(\sup_{z \in D} \|R_n(z)\| > \mathcal{K} \right) \rightarrow 0$$

as $n \rightarrow \infty$. Thus

$$Y_n(z) = \frac{\sqrt{n}}{z^2} \left[\frac{1}{n} \psi R_{n,l}(z) \phi \right] \mathbf{1}_{\Omega_{n,l}} + o(1)$$

in probability.

We now apply Lemma 7. Let \tilde{M}_n be the truncated matrix in Lemma 7 and define \tilde{X} , $\tilde{M}_{n,l}$, $\tilde{\phi}$, $\tilde{\psi}$ to be the appropriate blocks of \tilde{M}_n . Since M_n and \tilde{M}_n agree with probability going to 1, we have that

$$Y_n(z) = \frac{\sqrt{n}}{z^2} \left[\frac{1}{n} \tilde{\psi} R_{n,l}(z) \tilde{\phi} \right] \mathbf{1}_{\Omega_{n,l}} + o(1)$$

in probability.

Our goal is to apply Lemma 14 from Appendix A to the term $\tilde{\psi} R_{n,l}(z) \tilde{\phi}$. However, we first need to show that

$$\frac{1}{n} \text{tr} (R_{n,l}(z) R_{n,l}(z)^T) \Omega_{n,l}$$

converges in probability to a constant as n tends to infinity. We do so in the following lemma.

Lemma 12. *For any $z, w \in D$,*

$$\frac{1}{n} \text{tr} (R_n(z) R_n(w)^T) \mathbf{1}_{\Omega_n} \rightarrow \frac{1}{zw - 1}$$

in probability as $n \rightarrow \infty$. The result also holds when R_n is replaced by $R_{n,l}$ and Ω_n is replaced by $\Omega_{n,l}$.

Proof. We will prove the result for R_n on the event Ω_n . The proof for $R_{n,l}$ follows the same argument.

We define $X_n := \frac{1}{\sqrt{n}}M_n$ and recall that

$$\Omega_n = \{X_n : \|X_n\| \leq \kappa\}.$$

By Lemma 7, the matrix X_n and the truncated version \tilde{X}_n coincide with probability going to one. Thus it suffices to prove Lemma 12 for the truncated version. To avoid unnecessary notation, we let X_n denote the truncated matrix and $R_n(z)$ denote the resolvent of the truncated matrix.

Let $D(z) = (z - X_n)$ and $H(z, w) = [D(w)^T D(z)]^{-1}$. We will let $d_k(z)$ denote the k -th column of $D(z)$ and $D_k(z)$ denote the matrix $D(z)$ with the k -th column removed. Then we have that

$$\begin{aligned} \frac{1}{n} \text{tr} H(z, w) \mathbf{1}_{\Omega_n} &= \frac{1}{n} \sum_{k=1}^{n-1} \frac{1}{d_k(w)^T d_k(z) - d_k(w)^T D_k(z) H_k(z, w) D_k(w)^T d_k(z)} \mathbf{1}_{\Omega_n} \\ &= \frac{1}{n} \sum_{k=1}^{n-1} \frac{1}{d_k(w)^T d_k(z) - d_k(w)^T D_k(z) H_k(z, w) D_k(w)^T d_k(z) \mathbf{1}_{\Omega_n}} \mathbf{1}_{\Omega_n} \\ &= \frac{1}{n} \sum_{k=1}^{n-1} \beta_k. \end{aligned}$$

where $H_k(z, w) = [D_k(w)^T D_k(z)]^{-1}$ and $\beta_k^{-1} = d_k(w)^T d_k(z) - d_k(w)^T D_k(z) H_k(z, w) D_k(w)^T d_k(z) \mathbf{1}_{\Omega_n}$.

Let r_k denote the k -th column of X_n and let $X_n^{(k)}$ denote the matrix X_n with the k -th column removed. On the event Ω_n , we have that $\|X_n^{(k)}\| \leq \kappa$ and hence

$$\|H(z, w)\| \leq \mathcal{K}^2 \quad \text{and} \quad \sup_k \|H_k(z, w)\| \leq \mathcal{K}^2.$$

We also note that

$$d_k(w)^T d_k(z) = zw - (z + w)X_{nkk} + r_k^T r_k.$$

Recalling that the entries of X_n (and hence r_k) have variance $\frac{1}{n}$, we have that

$$\begin{aligned} \mathbb{P} \left(\sup_{1 \leq k \leq n-1} |d_k(w)^T d_k(z) - (zw + 1)| > \epsilon \right) &\leq \sum_{k=1}^{n-1} \left[\mathbb{P} \left(|X_{nkk}| > \frac{\epsilon}{2|z| + 2|w|} \right) + \mathbb{P} \left(|r_k^T r_k - 1| > \frac{\epsilon}{2} \right) \right] \\ &\leq \sum_{k=1}^{n-1} \left[\frac{(2|z| + 2|w|)^4 m_4}{n^2 \epsilon^4} + \frac{16}{\epsilon^4} \mathbb{E} |r_k^T r_k - 1|^4 \right] = O(\epsilon_n) \end{aligned}$$

by Lemma 10.

Define the event

$$\Omega_n^{(k)} = \{\|X_n\| : \|X_n^{(k)}\| \leq \kappa\}.$$

Below we will make use of the fact that $\mathbf{1}_{\Omega_n} \leq \mathbf{1}_{\Omega_n^{(k)}}$.

We note that

$$\begin{aligned} d_k(w)^T D_k(z) H_k(z, w) D_k(w)^T d_k(z) \\ = r_k^T D_k(z) H_k(z, w) D_k(w)^T r_k - w e_k^T D_k(z) H_k(z, w) D_k(w)^T r_k \\ - z r_k^T D_k(z) H_k(z, w) D_k(w)^T e_k + z w e_k^T D_k(z) H_k(z, w) D_k(w)^T e_k. \end{aligned}$$

Since $\|D_k(z)H_k(z, w)D_k(w)^T d_k(z)\| \mathbf{1}_{\Omega_n} \leq \kappa^2 \mathcal{K}^2$, by Lemma 10 and Markov's inequality, we have that

$$\begin{aligned} & \mathbb{P} \left(\sup_{1 \leq k \leq n-1} |r_k^T D_k(z) H_k(z, w) D_k(w)^T r_k - \frac{1}{n} \text{Tr}(D_k(z) H_k(z, w) D_k(w)^T)| \mathbf{1}_{\Omega_n} > \epsilon \right) \\ & \leq \mathbb{P} \left(\sup_{1 \leq k \leq n-1} |r_k^T D_k(z) H_k(z, w) D_k(w)^T r_k - \frac{1}{n} \text{Tr}(D_k(z) H_k(z, w) D_k(w)^T)| \mathbf{1}_{\Omega_n^{(k)}} > \epsilon \right) \\ & = O(\epsilon_n), \\ & \mathbb{P} \left(\sup_{1 \leq k \leq n-1} |e_k^T D_k(z) H_k(z, w) D_k(w)^T r_k| \mathbf{1}_{\Omega_n} > \epsilon \right) = O\left(\frac{1}{n}\right), \\ & \mathbb{P} \left(\sup_{1 \leq k \leq n-1} |r_k^T D_k(z) H_k(z, w) D_k(w)^T e_k| \mathbf{1}_{\Omega_n} > \epsilon \right) = O\left(\frac{1}{n}\right). \end{aligned}$$

Furthermore,

$$\frac{1}{n} \text{Tr}(D_k(z) H_k(z, w) D_k(w)^T) = \frac{n-2}{n} = 1 + O\left(\frac{1}{n}\right).$$

Therefore, the above estimates imply that

$$\sup_{1 \leq k \leq n-1} |\beta_k - zw + zwe_k^T D_k(z) H_k(z, w) D_k(w)^T e_k| \mathbf{1}_{\Omega_n} \longrightarrow 0$$

in probability as $n \rightarrow \infty$.

Let y_k denote the k -th row of X_n with the k -th entry removed. Then

$$e_k^T D_k(z) H_k(z, w) D_k(w)^T e_k = y_k^T H_k(z, w) y_k.$$

Let $\hat{H}_k(z, w) = [\hat{D}_k(w)^T \hat{D}_k(z)]^{-1}$ where $\hat{D}_k(z)$ is the matrix $D_k(z)$ with the k -th row removed. Since $H_k(z, w)^{-1} = \hat{H}_k(z, w)^{-1} + y_k y_k^T$, by the resolvent identity

$$y_k^T H_k(z, w) y_k = 1 - \frac{1}{1 + y_k^T \hat{H}_k(z, w) y_k}.$$

Since for any $1 \leq k \leq n-1$,

$$\begin{aligned} \left| \frac{1}{\beta_k} \right| &= |H(z, w)_{kk}| \mathbf{1}_{\Omega_n} \leq \mathcal{K}^2, \\ \left| y_k^T \hat{H}_k(z, w) y_k \right| \mathbf{1}_{\Omega_n} &\leq \kappa^2 \mathcal{K}^2, \end{aligned}$$

it follows that

$$\sup_{1 \leq k \leq n-1} \left| \frac{1}{\beta_k} - \frac{1 + y_k^T \hat{H}_k(z, w) y_k}{zw} \right| \mathbf{1}_{\Omega_n} \longrightarrow 0$$

in probability as $n \rightarrow \infty$.

Now define the event

$$\tilde{\Omega}_n^{(k)} = \left\{ X_n : \|\tilde{X}_n^{(k)}\| \leq \kappa \right\}$$

where $\tilde{X}_n^{(k)}$ is the matrix X_n with the k -th column and k -th row removed. Again we will make use of the fact that $\mathbf{1}_{\Omega_n} \leq \mathbf{1}_{\tilde{\Omega}_n^{(k)}}$. In particular, using Markov's inequality

and Lemma 10, it follows that

$$\begin{aligned} & \sup_{1 \leq k \leq n-1} \left| y_k^T \hat{H}_k(z, w) y_k - \frac{1}{n} \text{tr}(\hat{H}_k(z, w)) \right| \mathbf{1}_{\Omega_n} \\ & \leq \sup_{1 \leq k \leq n-1} \left| y_k^T \hat{H}_k(z, w) y_k - \frac{1}{n} \text{tr}(\hat{H}_k(z, w)) \right| \mathbf{1}_{\tilde{\Omega}_n^{(k)}} \longrightarrow 0 \end{aligned}$$

in probability as $n \rightarrow \infty$.

Finally, let $S_k(z, w) = [(w - \hat{X}_k)^T(z - \hat{X}_k)]^{-1}$ where \hat{X}_k is the matrix X_n where the k -th row and k -th column are replaced by zeros. In particular, $S_k(z, w)$ is an $n \times n$ matrix. It follows that

$$\frac{1}{n} \left| \text{tr}S_k(z, w) - \text{tr}\hat{H}_k(z, w) \right| = O\left(\frac{1}{n}\right).$$

Then by the resolvent identity, for $1 \leq k \leq n-1$,

$$\begin{aligned} & \left| \frac{1}{n} \text{tr}S_k(z, w) - \frac{1}{n} H(z, w) \right| \mathbf{1}_{\Omega_n} \\ & = \frac{1}{n} \text{tr} \left[H(z, w) \left((w - X_n)^T(z - X_n) - (w - \hat{X}_k)^T(z - \hat{X}_k) \right) S_k(z, w) \right] \mathbf{1}_{\Omega_n} \\ & \leq \frac{4}{n} \|H(z, w)\| \left(\|(w - X_n)^T(z - X_n)\| + \|(w - \hat{X}_k)^T(z - \hat{X}_k)\| \right) \|S_k(z, w)\| \mathbf{1}_{\Omega_n} \\ & \leq \frac{8}{n} \mathcal{K}^4 (|z| + \kappa) (|w| + \kappa) \end{aligned}$$

since $(w - X_n)^T(z - X_n) - (w - \hat{X}_k)^T(z - \hat{X}_k)$ is at most rank 4. Therefore, we have that

$$\sup_{1 \leq k \leq n-1} \left| \frac{1}{n} \text{tr}\hat{H}_k(z, w) - \frac{1}{n} H(z, w) \right| \mathbf{1}_{\Omega_n} = O\left(\frac{1}{n}\right).$$

Therefore, combining the above, yields that

$$\sup_{1 \leq k \leq n-1} \left| \frac{1}{\beta_k} - \frac{1 + \frac{1}{n} \text{tr}H(z, w)}{zw} \right| \mathbf{1}_{\Omega_n} \longrightarrow 0$$

in probability as $n \rightarrow \infty$ and hence

$$\frac{1}{n} \text{tr}H(z, w) \mathbf{1}_{\Omega_n} - \frac{1 + \frac{1}{n} \text{tr}H(z, w)}{zw} \mathbf{1}_{\Omega_n} \longrightarrow 0$$

in probability as $n \rightarrow \infty$. The proof of the lemma is complete. \square

By Lemma 12, we have that

$$\begin{aligned} & \frac{1}{n} \text{tr} \left(\frac{R_{n,l}(z)}{2z^2} + \frac{\overline{R_{n,l}(z)}}{2\bar{z}^2} \right) \left(\frac{R_{n,l}(w)}{2w^2} + \frac{\overline{R_{n,l}(w)}}{2\bar{w}^2} \right)^T \mathbf{1}_{\Omega_{n,l}} \\ & \quad \longrightarrow \frac{1}{4} \left[\frac{1}{z^2 w^2 (zw - 1)} + \frac{1}{z^2 \bar{w}^2 (z\bar{w} - 1)} + \frac{1}{\bar{z}^2 w^2 (\bar{z}w - 1)} + \frac{1}{\bar{z}^2 \bar{w}^2 (\bar{z}\bar{w} - 1)} \right], \\ & \frac{1}{n} \text{tr} \left(\frac{R_{n,l}(z)}{2\sqrt{-1}z^2} - \frac{\overline{R_{n,l}(z)}}{2\sqrt{-1}\bar{z}^2} \right) \left(\frac{R_{n,l}(w)}{2\sqrt{-1}w^2} - \frac{\overline{R_{n,l}(w)}}{2\sqrt{-1}\bar{w}^2} \right)^T \mathbf{1}_{\Omega_{n,l}} \\ & \quad \longrightarrow -\frac{1}{4} \left[\frac{1}{z^2 w^2 (zw - 1)} - \frac{1}{z^2 \bar{w}^2 (z\bar{w} - 1)} - \frac{1}{\bar{z}^2 w^2 (\bar{z}w - 1)} + \frac{1}{\bar{z}^2 \bar{w}^2 (\bar{z}\bar{w} - 1)} \right], \\ & \frac{1}{n} \text{tr} \left(\frac{R_{n,l}(z)}{2z^2} + \frac{\overline{R_{n,l}(z)}}{2\bar{z}^2} \right) \left(\frac{R_{n,l}(w)}{2\sqrt{-1}w^2} - \frac{\overline{R_{n,l}(w)}}{2\sqrt{-1}\bar{w}^2} \right)^T \mathbf{1}_{\Omega_{n,l}} \\ & \quad \longrightarrow \frac{1}{4\sqrt{-1}} \left[\frac{1}{z^2 w^2 (zw - 1)} - \frac{1}{z^2 \bar{w}^2 (z\bar{w} - 1)} + \frac{1}{\bar{z}^2 w^2 (\bar{z}w - 1)} - \frac{1}{\bar{z}^2 \bar{w}^2 (\bar{z}\bar{w} - 1)} \right], \end{aligned}$$

in probability as $n \rightarrow \infty$.

Fix $r \geq 1$, pick $z_1, \dots, z_r, w_1, \dots, w_r \in D$, and let $\alpha_k, \beta_k \in \mathbb{R}$ for $1 \leq k \leq r$. Then define the matrix

$$Q := \sum_{k=1}^r \left[\alpha_k \left(\frac{R_{n,l}(z_k)}{2z_k^2} + \frac{R_{n,l}(\bar{z}_k)}{2\bar{z}_k^2} \right) + \beta_k \left(\frac{R_{n,l}(w_k)}{2\sqrt{-1}w_k^2} - \frac{R_{n,l}(\bar{w}_k)}{2\sqrt{-1}\bar{w}_k^2} \right) \right] \mathbf{1}_{\Omega_{n,l}}.$$

Clearly $\tilde{\psi}$, $\tilde{\phi}$, and Q are independent. By applying Lemma 14 from Appendix A to the term

$$\frac{1}{\sqrt{n}} \tilde{\psi} Q \tilde{\phi}$$

and using the calculations above from Lemma 12, it follows that the finite dimensional distributions of $S_n(z)$ converge to $S(z)$.

4. TIGHTNESS

We now extend the finite dimensional convergence of $S_n(z)$ to weak convergence in the space of continuous functions by verifying that $\{S_n\}_{n \geq 1}$ is tight in the space of continuous functions on D (see for example [8]). Since $S_n(z)$ is expressed as a linear combination of $Y_n(z)_{ij}$, $1 \leq i, j \leq l$, it suffices to show that $\{\|Y_n\|\}_{n \geq 1}$ is tight in the space of continuous functions on D .

We remind the reader that

$$Y_n(z) = \sqrt{n} \left[\left(z - \frac{1}{\sqrt{n}} M_n \right)_{ij}^{-1} - \frac{1}{z} \delta_{ij} - \frac{m_{nij}}{z^2 \sqrt{n}} \right]_{i,j=1}^l.$$

Since D is compact, there exists some constant $C(D) > 0$ such that

$$(8) \quad 2 + \epsilon \leq |z| \leq C(D) \text{ for all } z \in D.$$

By Proposition 11,

$$\mathbb{P} \left(\sup_{z \in D} \|Y_n(z)\| \mathbf{1}_{\Omega_n^C} > \eta \right) \leq \mathbb{P} \left(\sup_{z \in D} \|R_n(z)\| > \mathcal{K} \right) + \mathbb{P}(\mathbf{1}_{\Omega_n^C}) \longrightarrow 0$$

as $n \rightarrow \infty$. The problem then reduces to showing that $\{\|Y_n\| \mathbf{1}_{\Omega_n}\}_{n \geq 1}$ is tight in the space of continuous functions on D . We will verify the Arzela-Ascoli criteria ([8]) by checking that there exists constants $\alpha, \beta, C > 0$ such that

$$\mathbb{E}|\|Y_n(z)\| \mathbf{1}_{\Omega_n} - \|Y_n(w)\| \mathbf{1}_{\Omega_n}|^\alpha \leq C|z - w|^{1+\beta}$$

for all $z, w \in D$ and all $n \geq 1$.

Lemma 13. *There exists a constant $C > 0$ (depending on the set D) such that*

$$\mathbb{E}|\|Y_n(z)\| \mathbf{1}_{\Omega_n} - \|Y_n(w)\| \mathbf{1}_{\Omega_n}|^2 \leq C|z - w|^2$$

for all $z, w \in D$ and all $n \geq 1$.

Proof. Define the $l \times l$ matrix

$$Q_n(z) := \left[R_n(z)_{ij} - \frac{1}{z} \delta_{ij} \right]_{i,j=1}^l.$$

By the triangle inequality,

$$(9) \quad \begin{aligned} |\|Y_n(z)\| \mathbf{1}_{\Omega_n} - \|Y_n(w)\| \mathbf{1}_{\Omega_n}| &\leq \|Y_n(z) - Y_n(w)\| \mathbf{1}_{\Omega_n} \\ &\leq \sqrt{n} \|Q_n(z) - Q_n(w)\| \mathbf{1}_{\Omega_n} + \left\| \frac{X}{z^2} - \frac{X}{w^2} \right\|. \end{aligned}$$

We start by bounding the second term on the right-hand side of (9). For $z, w \in D$,

$$\left| \frac{\|X\|}{z^2} - \frac{\|X\|}{w^2} \right|^2 \leq \frac{\|X\|^2}{|z|^4 |w|^4} |w^2 - z^2|^2 \leq \frac{\|X\|^2}{(2+\epsilon)^8} 4C(D)^2 |w - z|^2.$$

We then apply the naive bound

$$(10) \quad \mathbb{E}\|X\|^2 \leq \mathbb{E}\text{tr}(XX^*) = \sum_{i,j=1}^l \mathbb{E}|m_{nij}|^2 = l^2$$

to obtain

$$\mathbb{E} \left| \frac{\|X\|}{z^2} - \frac{\|X\|}{w^2} \right|^2 \leq C|z - w|^2$$

for some constant $C > 0$ (which depends on the set D).

In order to deal with the first term on the right-hand side of (9), we decompose

$$\begin{aligned} \sqrt{n} \left[R_n(z)_{ij} - \frac{1}{z} \delta_{ij} \right]_{i,j=1}^l &= \sqrt{n} \left(\left(z - \frac{1}{\sqrt{n}} X - \frac{1}{n} \psi R_{n,l}(z) \phi \right)^{-1} - \frac{1}{z} I_l \right) \\ &= -\frac{1}{z} \left(z - \frac{1}{\sqrt{n}} X - \frac{1}{n} \psi R_{n,l}(z) \phi \right)^{-1} \left(X + \frac{1}{\sqrt{n}} \psi R_{n,l}(z) \phi \right) \\ &= -\frac{1}{z} T_n(z) \left(X + \frac{1}{\sqrt{n}} \psi R_{n,l}(z) \phi \right) \end{aligned}$$

by the resolvent identity, where

$$T_n(z) = \left(z - \frac{1}{\sqrt{n}} X - \frac{1}{n} \psi R_{n,l}(z) \phi \right)^{-1}.$$

We note that all the relevant inverses above exist on the event Ω_n . Then

$$\sqrt{n} \|Q_n(z) - Q_n(w)\| \leq \left\| \frac{X}{w} T_n(w) - \frac{X}{z} T_n(z) \right\| + \left\| \frac{T_n(w)}{w\sqrt{n}} \psi R_{n,l}(w) \phi - \frac{T_n(z)}{z\sqrt{n}} \psi R_{n,l}(z) \phi \right\|$$

On the event Ω_n , we have that

$$\begin{aligned} \left\| \frac{X}{w} T_n(w) - \frac{X}{z} T_n(z) \right\| &\leq \|X\| \left(\|T_n(w)\| \left| \frac{1}{w} - \frac{1}{z} \right| + \frac{1}{|z|} \|T_n(w) - T_n(z)\| \right) \\ &\leq \|X\| \left(\frac{\mathcal{K}}{(2+\epsilon)^2} |z-w| + \|T_n(w) - T_n(z)\| \right) \end{aligned}$$

and by the resolvent identity

$$(11) \quad \|T_n(w) - T_n(z)\| \leq \mathcal{K}^2 |z-w|.$$

Thus by (10), we obtain that

$$\mathbb{E} \left\| \frac{X}{w} T_n(w) - \frac{X}{z} T_n(z) \right\|^2 \mathbf{1}_{\Omega_n} \leq C |z-w|^2.$$

Similarly, we now bound

$$\left\| \frac{T_n(w)}{w\sqrt{n}} \psi R_{n,l}(w) \phi - \frac{T_n(z)}{z\sqrt{n}} \psi R_{n,l}(z) \phi \right\|.$$

Using the triangle inequality and (11), it suffices to bound

$$(12) \quad \frac{1}{n} \mathbb{E} \|\psi R_{n,l}(z) \phi\|^2 \mathbf{1}_{\Omega_n}$$

and

$$(13) \quad \frac{1}{n} \mathbb{E} \|\psi [R_{n,l}(z) - R_{n,l}(w)] \phi\|^2 \mathbf{1}_{\Omega_n}.$$

For (12), we note that

$$\begin{aligned} \frac{1}{n} \mathbb{E} \|\psi R_{n,l}(z) \phi\|^2 \mathbf{1}_{\Omega_n} &\leq \frac{1}{n} \mathbb{E} \|\psi R_{n,l}(z) \phi\|^2 \mathbf{1}_{\Omega_{n,l}} \\ &\leq \frac{1}{n} \sum_{i,j=1}^l \mathbb{E} |(\psi R_{n,l}(z) \phi)_{ij}|^2 \leq l^2 \mathcal{K}^2 \end{aligned}$$

by Lemma 9.

For (13), we apply the resolvent identity and obtain

$$\begin{aligned} \frac{1}{n} \mathbb{E} \|\psi [R_{n,l}(z) - R_{n,l}(w)] \phi\|^2 \mathbf{1}_{\Omega_n} &\leq \frac{1}{n} |z-w|^2 \mathbb{E} \|\psi R_{n,l}(z) R_{n,l}(w) \phi\|^2 \mathbf{1}_{\Omega_n} \\ &\leq \frac{1}{n} |z-w|^2 \mathbb{E} \|\psi R_{n,l}(z) R_{n,l}(w) \phi\|^2 \mathbf{1}_{\Omega_{n,l}}. \end{aligned}$$

Since $\psi R_{n,l}(z) R_{n,l}(w) \phi$ is an $l \times l$ matrix, we again apply Lemma 9 to each entry to bound the norm and obtain

$$\frac{1}{n} \mathbb{E} \|\psi R_{n,l}(z) R_{n,l}(w) \phi\|^2 \mathbf{1}_{\Omega_{n,l}} \leq l^2 \mathcal{K}^4.$$

Therefore

$$\mathbb{E} \left\| \frac{T_n(w)}{w\sqrt{n}} \psi R_{n,l}(w) \phi - \frac{T_n(z)}{z\sqrt{n}} \psi R_{n,l}(z) \phi \right\|^2 \mathbf{1}_{\Omega_n} \leq C |z-w|^2$$

and the proof of the lemma is complete. \square

Acknowledgment. The author would like to thank an anonymous referee for the valuable suggestions and remarks.

APPENDIX A. CENTRAL LIMIT THEOREM

Lemma 14 below is a Corollary of Theorem A.5 from [19]. The version presented here is for truncated random variables.

Lemma 14. *Let $\{B^s : 1 \leq s \leq r\}$ be a family of $n \times n$ real random matrices. Let $\{\mathbf{x}_n^{(s)}, \mathbf{y}_n^{(s)} : 1 \leq s \leq r\}$ be a collection of independent n -vectors with independent real standardized entries where $\mathbf{y}_n^{(s)} = (y_{nj}^{(s)})_{1 \leq j \leq n}$, $\mathbf{x}_n^{(s)} = (x_{nj}^{(s)})_{1 \leq j \leq n}$, and*

$$\sup_{n,j} \left(\mathbb{E}|y_{nj}^{(s)}|^4 + \mathbb{E}|x_{nj}^{(s)}|^4 \right) \leq m_4 < \infty$$

for $1 \leq s \leq r$. Further assume that

- (i) for each $1 \leq s \leq r$, $B^s, \mathbf{x}_n^{(s)}$, and $\mathbf{y}_n^{(s)}$ are independent,
- (ii) there exists a sequence $\epsilon_n \rightarrow 0$ such that $\sup_{n,j,s} |x_{nj}^{(s)}| \leq \epsilon_n n^{1/4}$ and $\sup_{n,j,s} |y_{nj}^{(s)}| \leq \epsilon_n n^{1/4}$,
- (iii) there exists a constant a (not depending on n) such that $\max_{1 \leq s \leq r} \|B^s\| \leq a$,
- (iv) for each $1 \leq s \leq r$, $\frac{1}{n} \text{tr}((B^s)^* B^s)$ converges in probability to a number $a_2(s)$.

Then the random r -vector

$$\mathbf{Z}_n = \frac{1}{\sqrt{n}} \left((\mathbf{x}_n^{(s)})^T B^s \mathbf{y}_n^{(s)} \right)_{1 \leq s \leq r}$$

converges in distribution as $n \rightarrow \infty$ to an r -vector $\mathbf{Z} = (z_s)_{1 \leq s \leq r}$ whose entries are independent normal random variables with mean zero and where the variance of z_s is $a_2(s)$.

REFERENCES

- [1] Anderson G.W., Guionnet A., and Zeitouni O. *An Introduction to Random Matrices*, Cambridge Studies in Advanced Mathematics 118, Cambridge University Press, New York, 2010.
- [2] Z. D. Bai, *Circular law*, Ann. Probab. 25 (1997), 494-529.
- [3] Bai, Z. D. *Methodologies in spectral analysis of large-dimensional random matrices, a review*. Statist. Sinica 9, 611–677 (1999).
- [4] Z. D. Bai, J. Silverstein, Spectral analysis of large dimensional random matrices, Mathematics Monograph Series 2, Science Press, Beijing 2006.
- [5] Bai, Z.D. and Silverstein, J. W., *CLT for linear spectral statistics of large-dimensional sample covariance matrices*, Ann. Probab. 32, 533-605 (2004).
- [6] Bai, Z. D., Pan, G. M., *Limiting Behavior of Eigenvectors of Large Wigner Matrices*, Journal of Statistical Physics, Volume 146, Number 3 (2012), 519-549.
- [7] Billingsley P., *Probability and Measure*, 3rd edition, Wiley Series in Probability and Mathematical Statistics. Wiley, New York, 1995.
- [8] Billingsley P., *Convergence of Probability Measures*, Wiley, New York, 1968.
- [9] Durrett R., *Probability. Theory and Examples*, 4th ed., Cambridge University Press, New York, 2010.
- [10] A. Edelman, *The Probability that a Random Real Gaussian Matrix has k Real Eigenvalues, Related Distributions, and the Circular Law*, J. Multivariate Anal. 60, 203–232, (1997).
- [11] J. Ginibre, *Statistical ensembles of complex, quaternion and real matrices*, J. Math. Phys. 6 (1965), 440–449.
- [12] V. L. Girko, *Circular law*, Theory Probab. Appl. (1984), 694-706.
- [13] V. L. Girko, *The strong circular law, Twenty years later*, II. Random Oper. Stochastic Equations 12 (2004), no. 3, 255-312.
- [14] F. Götze and A. N. Tikhomirov, *The circular law for random matrices*, Ann. Prob., 38 (2010), no. 4, 1444-1491.
- [15] Horn, R., Johnson, C., *Matrix analysis*, Cambridge University Press, Cambridge; New York (1985)

- [16] Lytova A. and Pastur L., *Fluctuations of matrix elements of regular functions of Gaussian random matrices*, J. Stat. Phys., 134, 147-159 (2009).
- [17] M. L. Mehta, *Random matrices and the statistical theory of energy levels*, Acad. Press. (1967).
- [18] Nourdin, Ivan and Peccati, Giovanni, *Universal Gaussian fluctuations of non-Hermitian matrix ensembles: from weak convergence to almost sure CLTs*, ALEA Lat. Am. J. Probab. Math. Stat. 7 (2010), 341375.
- [19] S. O'Rourke, D. Renfrew, A. Soshnikov, *On fluctuations of matrix entries of regular functions of Wigner matrices with non-identically distributed entries*, arXiv:1104.1663v2 [math.PR]
- [20] S. O'Rourke, D. Renfrew, A. Soshnikov, *Fluctuations of Matrix Entries of Regular Functions of Sample Covariance Random Matrices*, arXiv:1106.0320v1 [math.PR].
- [21] G. Pan and W. Zhou, *Circular law, extreme singular values and potential theory*, Journal of Multivariate Analysis, 101 (2010), 645-656.
- [22] Pastur L. and Lytova A., *Non-Gaussian limiting laws for the entries of regular functions of the Wigner matrices*, available at arXiv:1103.2345 [math.PR].
- [23] Pizzo, A., Renfrew D., and Soshnikov A., *Fluctuations of matrix entries of regular functions of Wigner matrices*, available at arXiv:1103.1170 [math.PR].
- [24] Pizzo, A., Renfrew D., and Soshnikov A., *On finite rank deformations of Wigner matrices*, available at arXiv:1103.3731 [math.PR].
- [25] Rider, B. and Silverstein, Jack W., *Gaussian fluctuations for non-Hermitian random matrix ensembles*, Ann. Probab. 34 (2006), no. 6, 21182143.
- [26] T. Tao and V. Vu, *Random matrices: the circular law*, Communication in Contemporary Mathematics 10 (2008), 261-307.
- [27] T. Tao and V. Vu, *Random matrices: universality of ESDs and the circular law*, Ann. Probab. 38 (2010), no. 5p. 2023–2065, With an appendix by Manjunath Krishnapur.

DEPARTMENT OF MATHEMATICS, RUTGERS, PISCATAWAY, NJ 08854

E-mail address: sdo21@math.rutgers.edu